

Sequence listing

5 <110> GlaxosmithKline Biologicals sa
Glaxo Group Ltd

<120> Immunogenic compositions

10 <130> B45310

<160> 11

15 <170> FastSEQ for Windows Version 4.0

<210> 1
<211> 553
<212> PRT

20 <213> Rattus norvegicus

<400> 1
Met Ile Gln Arg Leu Trp Ala Ser Arg Leu Leu Arg His Arg Lys Ala
1 5 10 15
25 Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
20 25 30
Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
35 40 45
Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
50 55 60
30 Leu Val Ser Val Pro Leu Leu Gly Ser Ala Ser Asp Gln Trp Arg Gly
65 70 75 80
Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Val
85 90 95
35 Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu
100 105 110
Leu Cys Ser Asp Thr Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly
115 120 125
40 Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu
130 135 140
Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala
145 150 155 160
Phe Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr
165 170 175
45 Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu
180 185 190
Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
195 200 205
Ile Cys Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Val Leu Gly
210 215 220
50 Pro Pro Glu Pro Ala Glu Gly Leu Leu Val Ser Ser Val Ser Arg Arg
225 230 235 240
Cys Cys Ser Cys His Ala Gly Leu Ala Phe Arg Asn Leu Gly Thr Leu
245 250 255
55 Phe Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
260 265 270
Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
275 280 285

Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
 290 295 300
 Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
 305 310 315 320
 5 Ile Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
 325 330 335
 Phe Phe Ser Leu Val Met Asp Arg Leu Val Gln Lys Phe Gly Thr Arg
 340 345 350
 10 Ser Val Tyr Leu Ala Ser Val Met Thr Phe Pro Val Ala Ala Ala Ala
 355 360 365
 Thr Cys Leu Ser His Ser Val Val Val Val Thr Ala Ser Ala Ala Leu
 370 375 380
 Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
 385 390 395 400
 15 Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
 405 410 415
 Asp Ala Gly Gly Gly Ser Ser Glu Asp Ser Gln Thr Thr Ser Phe Leu
 420 425 430
 20 Leu Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Pro
 435 440 445
 Gly Gly Ser Ser Ile Leu Val Pro Pro Pro Ala Leu Cys Gly Ala Ser
 450 455 460
 Ala Cys Asp Val Ser Met Arg Val Val Val Gly Glu Pro Pro Glu Ala
 465 470 475 480
 25 Lys Val Val Thr Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 485 490 495
 Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
 500 505 510
 30 Ile Val Gln Leu Ser His Ser Val Thr Ala Tyr Met Val Ser Ala Ala
 515 520 525
 Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
 530 535 540
 Lys Asn Asp Leu Ala Lys Tyr Ser Leu
 545 550

35

<210> 2

<211> 2314

<212> DNA

40

<213> Rattus norvegicus

<400> 2

gggctcttag acaccgcaac aaaagcaact ttcttccaag ccactgccac ctgttggggtt 60
 ttcacacatt tcgagcttta gttccgatcc ccagaacatc cacgtagttt ttctggcctt 120
 45 ctggctgagc catggaggcc gacagaggag gggagaagt tgaagcttga gaaggatttc 180
 cgtatgcgca aggctaccca tgcttgtcct tcttcccatg accctgggtca gccctcctct 240
 gccctcctct tcttgcccc cttctctcca ggggtccgact gacgagatgt gtccccatca 300
 agcaaggcac tagatggtga cgtgttcagt gtgggatgag atgccgaagt ggtactcaag 360
 ggctggccga aatgggagcc tggctgcacc ctccggaggtt ggtgctagca aggaggagaa 420
 50 gccgcggcag ggctgactca aaacagctgt ggggtgtgtg aatggcccc ggaccctaa 480
 ccgccctgtc catcatgatc cagaggctgt gggccagccg tctgctaagg catcggaaag 540
 ccagctcct gctggtcaac ctgctaacct tcggcctgga ggtgtgcctg gctgctggca 600
 ttacctatgt gccaccctt ctgctggaag tcggggtaga ggaaaagttc atgacctagg 660
 tggtgggcat tggcccagtg ctgggcctgg tttctgttcc actcctaggc tcagccagtg 720
 55 accagtggcg tgggcgctat ggccgcggga gaccctttat ctgggctctg tccctgggtg 780
 tcttgctaag cctcttcctc atcccgaggg ccggtctggct ggcagggcta ctgtgttcag 840
 atactaggcc cctggagttg gccctgctca tcttgggagt ggggctgctg gacttttgcg 900
 gccaggtgtg ctttactcca ctggaggcct tactctccga cctcttccgg gaccagacc 960
 actgccgcca agccttctct gtctatgcct tcatgatcag cctcgggggc tgccctgggct 1020

acctcttacc tgccattgac tgggacacca gcgccttggc cccctaccta ggcactcagg
 1080
 aagaatgcct ctteggcctc ctcacctca tttttctcat ctgtgtggca gccactctgc
 1140
 5 ttgtggctga ggaggcagtc cttggccac cccagccagc agaaggggtg ttggtctcct
 1200
 ccgtgtcacg ccggtgctgc tcatgccatg ctggcctggc tttccggaat ctgggtaccc
 1260
 10 tgtttccccg gctgcaccag ctgtgctgcc gaatgcctcg caccctgcgc cggctctttg
 1320
 tggctgagct gtgcagctgg atggcactta tgactttcac actgttctac acggacttcg
 1380
 tgggagaggg gctgtaccag ggtgtcccca gagcagagcc aggtaccgag gcccgagagc
 1440
 15 actatgatga aggcattcga atgggcagcc tggggctctt cctgcagtgt gccatctccc
 1500
 tgttcttctc cctgggtcatg gacaggtgg tacagaagtt cggcacacgg tcaagtctacc
 1560
 tggccagtgt gatgacctt cccgtggctg ccgctgccac gtgcctgtcc cacagcgtgg
 1620
 20 ttgtagtac agcctcagct gccctcaccg ggttcacctt ctcagccttg cagatcctgc
 1680
 cttacacgct cgctccctc taccatcgag agaagcaggt gttcctgccc aaataccgag
 1740
 25 gggacgctgg aggtggtagc agtgaagaca gccaaacaac cagcttcttg ctaggcctta
 1800
 agccaggagc tcccttcccc aatggacacg tgggccttg cggcagcagc atcctggtgc
 1860
 30 cccacctgc actctgtggg gcctctgcct gtgatgtctc catgcgagtg gtagtgggtg
 1920
 agccacctga agccaaggtt gttactggac ggggcatttg cctggacctt gccatcctgg
 1980
 acagtgcctt tctgctgtcc caggtggctc cgtccctgtt catgggctcc attgtccagc
 2040
 35 tgagccactc tgtcactgcc tatatggtat cagctgcagg cttgggtctg gtcgccattt
 2100
 actttgtac acaggtagt tttgacaaga atgacttggc caaatactca ctgtagaatt
 2160
 40 ctgtaaggca tcaaagaaga ggatctgcct ccccggttct cagccccaga gggctgcaga
 2220
 gctggtctct ttcgggtctc tgttgccctg agtggctctc cactgccatc cgaaggcagt
 2280
 gaggtgtatg gctgcacagg ttggagcttt tggc
 2314
 45
 <210> 3
 <211> 553
 <212> PRT
 <213> Maccaca fascicularis
 50
 <400> 3
 Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
 1 5 10 15
 55 Gln Leu Leu Leu Ile Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
 20 25 30
 Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
 35 40 45
 Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
 50 55 60
 60 Leu Val Ser Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly

	65	Arg	Tyr	Gly	Arg	Arg	Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu	Gly	Ile	80
					85						90					95		
5		Leu	Leu	Ser	Leu	Phe	Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala	Gly	Leu	
					100					105						110		
		Leu	Cys	Pro	Asp	Pro	Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile	Leu	Gly	
					115					120					125			
		Val	Gly	Leu	Leu	Asp	Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro	Leu	Glu	
					130				135				140					
10		Ala	Leu	Leu	Ser	Asp	Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg	Gln	Ala	
							150					155					160	
		Tyr	Ser	Val	Tyr	Ala	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu	Gly	Tyr	
						165					170					175		
15		Leu	Leu	Pro	Ala	Ile	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro	Tyr	Leu	
					180					185					190			
		Gly	Thr	Gln	Glu	Glu	Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile	Phe	Leu	
					195					200				205				
		Thr	Cys	Val	Ala	Ala	Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala	Leu	Gly	
								215					220					
20		Pro	Ala	Glu	Pro	Ala	Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Pro	Ser	His	
							230					235					240	
		Cys	Cys	Pro	Cys	Trp	Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly	Ala	Leu	
						245					250					255		
25		Leu	Pro	Arg	Leu	His	Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr	Leu	Arg	
					260					265					270			
		Arg	Leu	Phe	Val	Ala	Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met	Thr	Phe	
					275				280					285				
		Thr	Leu	Phe	Tyr	Thr	Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln	Gly	Val	
							295						300					
30		Pro	Arg	Ala	Glu	Leu	Gly	Thr	Glu	Ala	Arg	Arg	His	Tyr	Asp	Glu	Gly	
							310					315					320	
		Val	Arg	Met	Gly	Ser	Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile	Ser	Leu	
						325					330					335		
35		Val	Phe	Ser	Leu	Val	Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly	Thr	Arg	
					340					345					350			
		Ala	Val	Tyr	Leu	Ala	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala	Gly	Ala	
					355					360					365			
		Thr	Cys	Leu	Ser	His	Ser	Val	Ala	Val	Val	Thr	Ala	Ser	Ala	Ala	Leu	
								375					380					
40																		

	545		550				
		<210> 4					
5		<211> 3514					
		<212> DNA					
		<213> <i>Maccaca fascicularis</i>					
		<400> 4					
10	aaaaaaaaag	ccgccggctg	gcgcgcgtgg	ggggcaagga	aaagaggggg	gaaaccagtc	60
	tgcacgcgct	ggctccgggt	gacagcccg	cgccataggcc	aggcagcgtc	tcctctctgtc	120
	accagactg	gaggcaatgt	tctgatact	gcacactgca	cccttgacct	cccagactca	180
	agcaatcttc	ccatctcagc	ctcttaagta	gctgggacca	caggatctga	gtgatgagat	240
	gtgtccccac	tgaggtgccc	cacagcagca	ggtgttgagc	atgggctgag	aagctggacc	300
15	ggcaccaaag	ggctggcgga	aatgggcgcc	tggctgattc	ctaggcagtt	ggcgccagca	360
	aggaggagag	gccgtggctt	ccggagcaga	gcggagacga	agcagttctg	gagtgtctaa	420
	acggccccct	gagccctacc	cgccctggccc	actatggctc	agaggctgtg	ggtgagccgc	480
	ctgctgcggc	accggaagc	ccagctcttg	ttgatcaacc	tgctaacctt	tggcctggaa	540
	gtgtgttttg	ccgcaggcat	cacctatgtg	ccgcctctgc	tgctggaagt	gggggtagaa	600
20	gagaagttca	tgacatgggt	gctgggcac	ggtccagtgc	tgggcctggg	ctctgtccca	660
	ctcctaggct	cagccagtga	ccactggcgc	ggacgctatg	gccgcggcg	gcccttcac	720
	tgggcgctgt	ccttgggcat	cctgctgagc	ctctttctca	tcccaagggc	tggctggctg	780
	gcagggtgct	tgtgcccga	tcccaggccc	ctggagctgg	cactgctcat	cctgggcgta	840
	gggctgctgc	acttctgtgg	ccagggtgtg	ttcactccac	tggaggccct	gctctctgac	900
25	ctgttccggg	accgggacca	ctgtcgccag	gcctactccg	tctatgcctt	catgatcagt	960
	cttgggggct	gcctgggcta	cctcctgcct	gccattgact	gggacaccag	tgccctggcc	1020
	ccctacctgg	gcacccagga	ggagtgcctc	tttggcctgc	tcacctcat	cttccctacc	1080
	tgcgtagcag	ccacactgct	ggtggccgag	gaggcagcac	tgggccccgc	cgagccagcg	1140
	gaagggctgt	ctgccccctc	cctgcccgtc	cactgctgtc	cgtgctgggc	ccgcctggct	1200
30	ttccggaacc	tgggcgcctt	gcttccccgg	ctgcaccage	tgtgctgccg	catgccccgc	1260
	accctgcgcc	ggctcttctg	ggctgagctg	tgcagctgga	tggcactcat	gaccttcacg	1320
	ctgtttttaca	oggatttctg	gggcgagggg	ctataccagg	gcgtgcccag	agctgagctg	1380
	ggcaccgagg	cccggagaca	ctatgatgaa	ggcggttcgga	tgggcagctc	ggggctgttc	1440
	ctgcagtgcg	ccatctccct	ggtcttctct	ctgggtcatg	accggctggg	gcagcgattc	1500
35	ggcactcgag	cagtctatct	ggccagtgtg	gcagctttcc	ctgtggctgc	cggtgccacg	1560
	tgctgttccc	acagtgtggc	tgtgggtgac	gcttcagccg	ccctcactgg	gttcaccttc	1620
	tcagccctgc	agatcctgcc	ctacacactg	gcctccctct	accaccggga	gaggcagggtg	1680
	ttcctgcccc	aataccgagg	ggacgctgga	ggcactagca	gtgaggacag	cctgatgact	1740
	agcttccctg	caggccctaa	gcctggagct	cccttcccta	atggacacgt	gggtgctgga	1800
40	ggcagtggcc	tgcttccacc	tccaccgcg	ctctgcgggg	cctctgcctg	cgatgtctct	1860
	gtacgtgtgg	tgggtgggtga	gcccaccgag	gccagggtgg	ttccggggcg	gggcatctgc	1920
	ctggacctcg	ccatcctgga	tagtgccctc	ctgctgtccc	aggtggcccc	gtccctgttc	1980
	atgggctcca	tgcgtccagct	cagccagctc	gtcactgcct	atatggtgtc	tgtgtcaggc	2040
	ctgggtctgg	ttgccattta	ctttgtctaca	caggtagtat	ttgacaagag	cgacttggcc	2100
45	aaatactcgg	tgtagaaaac	ttccagcaca	ttgggggtgga	gggcctgcct	cactgggtcc	2160
	cagctcccca	ctctttgtta	gccccatggg	gctgctgggc	tggccgccag	tttctgttgc	2220
	tgccaaagta	atgtggctct	ctgctgccac	cctgtgctgc	tgaggtgctg	agctgcacag	2280
	ctgggggctg	gggcatccct	ctccctcctc	cccagctctc	agggctgcct	gactggaagc	2340
	cttccaaggg	ggtttcagtc	tggacttcta	cagggaggct	agaagggcag	ggcatttgat	2400
50	tcgctccatg	cactggaatg	tggggactct	gcagggtggat	taccagggct	cagggttaac	2460
	agctagcctc	ctggctgaga	catacctaga	gaaggggttt	tgggagctga	gtaaactcag	2520
	tcacctgggt	tcccacctct	aagccccctt	aacctgcagc	ttcatttaat	gtagctcttg	2580
	catgggagtt	tctaggatga	aacactcctc	cgtgggattt	gaacgtatga	aagttatttg	2640
	taggggaaga	gtcctgaggg	gcaacacacc	aggtcccttc	agcccacagc	actgcctttt	2700
55	tgctgatccc	ctgactctta	ccttttatca	ggacgtggcc	tattgggtccc	tttgttgcca	2760
	tcatagggac	acaggcattt	aaatatattaa	cttattttatt	taacaaagta	gaagggaatc	2820
	catttgctagc	ttttgtgtgt	tgggtgtctaa	gatttgggta	gggtgggatc	cccaacaatc	2880
	aggtccactg	agatcactgg	tcattggggt	gatcattgac	agaatcttct	tctcctgggg	2940
	tctggctcct	caaaatgcct	aaccaggac	cttggaaatt	ttactcatcc	cgactgataa	3000
60	ttccaaatgc	tgttacccaa	ggttaggggg	ttgaagggaag	gtggaggggtg	gggcttcagg	3060

tctcaacagc ttccctaacc accccttttc tcttggccca gcctgggtcc cccacttct 3120
 actccctct actgtctcta ggactgggct gatgaaggca ctgcctgaaa tttccctcac 3180
 ccccaacttt cccactggc tccacaaccc tgtttggagc tggtgcagga ccagaagcac 3240
 aaagtgtggt ttcccaggcc tttgtccatc tcagccccc agagtatatc tgtgcttggg 3300
 5 gaatctcaca cagaaactca ggagcaccgc ctgcctgagc taaggaggct ttatctctca 3360
 ggggggttta agtgccgttt gcaataatgt cttatttatt tagcggggca aatattttat 3420
 actgtaagtg agcaatcagt ataatgttta tggatgatgaa attaaaggct ttcttatatg 3480
 tttaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 3514

10 <210> 5
 <211> 553
 <212> PRT
 <213> Homo sapiens

15 <400> 5
 Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
 1 5 10 15
 Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
 20 20 25 30
 Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
 35 40 45
 Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
 50 55 60
 Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly
 25 65 70 75 80
 Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile
 85 90 95
 Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu
 100 105 110
 30 Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly
 115 120 125
 Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu
 130 135 140
 Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala
 35 145 150 155 160
 Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr
 165 170 175
 Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu
 180 185 190
 40 Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
 195 200 205
 Thr Cys Val Ala Ala Thr Leu Val Ala Glu Glu Ala Ala Leu Gly
 210 215 220
 Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
 45 225 230 235 240
 Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu
 245 250 255
 Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
 260 265 270
 50 Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
 275 280 285
 Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
 290 295 300
 Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
 55 305 310 315 320
 Val Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
 325 330 335
 Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg
 340 345 350
 60 Ala Val Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala

355 360 365
 Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 370 375 380
 5 Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
 385 390 395 400
 Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
 405 410 415
 Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu
 420 425 430
 10 Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala
 435 440 445
 Gly Gly Ser Gly Leu Leu Pro Pro Pro Ala Leu Cys Gly Ala Ser
 450 455 460
 Ala Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
 465 470 475 480
 15 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 485 490 495
 Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
 500 505 510
 20 Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala
 515 520 525
 Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
 530 535 540
 Lys Ser Asp Leu Ala Lys Tyr Ser Ala
 545 550
 25

 <210> 6
 <211> 255
 30 <212> PRT
 <213> Homo sapiens

 <400> 6
 35 Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg
 1 5 10 15
 Arg His Tyr Asp Glu Gly Val Arg Met Gly Ser Leu Gly Leu Phe Leu
 20 25 30
 Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val Met Asp Arg Leu Val
 35 40 45
 40 Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala Ser Val Ala Ala Phe
 50 55 60
 Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His Ser Val Ala Val Val
 65 70 75 80
 45 Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile
 85 90 95
 Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg Glu Lys Gln Val Phe
 100 105 110
 Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser
 115 120 125
 50 Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro
 130 135 140
 Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu Leu Pro Pro Pro Pro
 145 150 155 160
 Ala Leu Cys Gly Ala Ser Ala Cys Asp Val Ser Val Arg Val Val Val
 165 170 175
 55 Gly Glu Pro Thr Glu Ala Arg Val Val Pro Gly Arg Gly Ile Cys Leu
 180 185 190
 Asp Leu Ala Ile Leu Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro
 195 200 205
 60 Ser Leu Phe Met Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala

210 215 220
 Tyr Met Val Ser Ala Ala Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala
 225 230 235 240
 Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala
 5 245 250 255

<210> 7
 <211> 231
 10 <212> PRT
 <213> Homo sapiens

<400> 7
 15 Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe
 1 5 10 15
 Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val
 20 20 25 30
 Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys
 35 40 45
 20 Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Leu Thr Gly
 50 55 60
 Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu
 65 70 75 80
 Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr
 25 85 90 95
 Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly
 100 105 110
 Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly
 115 120 125
 30 Ser Gly Leu Leu Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys
 130 135 140
 Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val
 145 150 155 160
 Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp Ser Ala
 35 165 170 175
 Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser Ile Val
 180 185 190
 Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala Ala Leu
 195 200 205
 40 Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp Lys Ser
 210 215 220
 Asp Leu Ala Lys Tyr Ser Ala
 225 230

45 <210> 8
 <211> 1788
 <212> DNA
 <213> Artificial Sequence

50 <220>
 <223> DNA sequence for alphaprepro signal sequence fused
 to human P501S and fused to a His tag

55 <400> 8
 atgagtttcc tcaattttac tgcagtttta ttgcgagcat cctccgcatt agctgctcca 60
 gtcaacacta caacagaaga tgaaacggca caaattccgg ctgaagctgt catcggttac 120
 tcagatttag aaggggattt cgatgttgct gttttgccat tttccaacag cacaataaac 180
 gggttattgt ttataaatat tactattgcc agcattgctg ctaaagaaga aggggtatct 240
 60 ctcgagaaaa gagaggctga agccatggtg ctggggcattg gtccagtgcg gggcctggtc 300

5 tgtgtcccgc tcctaggctc agccagtgc cactggcgtg gacgctatgg ccgcccgcgg 360
 cccttcacat gggcactgtc cttgggcac cgtctgagcc tctttctcat cccaagggcc 420
 ggctggctag cagggctgct gtgcccggat cccagggccc tggagctggc actgctcatc 480
 ctgggcgtgg ggctgctgga cttctgtggc caggtgtgct tcactccact ggaggccctg 540
 ctctctgacc tcttccggga cccggaccac tgtgccagg cctactctgt ctatgccttc 600
 atgatcagtc ttgggggctg cctgggctac ctctgcctg ccattgactg ggacaccagt 660
 gccctggccc cctacctggg caccaggag gagtgcctct ttggcctgct caccctcatc 720
 ttcctcacct gcgtagcagc cacactgctg gtggctgagg aggcagcgt gggccccacc 780
 gagccagcag aagggctgtc ggccccctcc ttgtcgcccc actgctgtcc atgccggggc 840
 10 cgcttggtt tccggaacct gggcgcctg cttccccggc tgcaccagct gtgctgccgc 900
 atgccccgca ccctgcgcgc gctcttcgtg gctgagctgt gcagctggat ggcactcatg 960
 accttcacgc tgttttacac ggatttcgtg ggcgagggg tgtaccagg cgtgcccaga 1020
 gctgagccgg gcaccgagg ccggagacac tatgatgaag gcgttcggat gggcagcctg 1080
 gggctgttcc tgcagtgcgc catctccctg gtcttctctc tggctcatgga ccggctgggtg 1140
 15 cagcgattcg gcactcgagc agtctatctg gccagtgtgg cagctttccc tgtggctgcc 1200
 ggtgccacat gcctgtccca cagtgtggcc gtggtgacag cttcagccgc cctcaccggg 1260
 ttcaccttct cagccctgca gatcctgccc tacacactgg cctccctcta ccaccgggag 1320
 aagcaggtgt tcctgcccac ataccgaggg gacactggag gtgctagcag tgaggacagc 1380
 ctgatgacca gcttccctg ccaggccctaag cctggagctc ccttccctaa tggacacgtg 1440
 20 ggtgctggag gcagtggcct gctccacact ccaccgcgc tctgcggggc ctctgcctgt 1500
 gatgtctccg tacgtgtggg ggtgggtgag cccaccgagg ccagggtggg tccgggcccgg 1560
 ggcactctgc tggacctgc catcctggat agtgccttcc tgcgtgtcca ggtgtcccca 1620
 tccctgttta tgggtccat tgtccagctc agccagtctg tcactgccta tatggtgtct 1680
 gccgcaggcc tgggtctggg cgccatttac tttgctacac aggtagtatt tgacaagagc 1740
 25 gacttggcca aatactcagc ggggtggacac catcaccatc accattaa 1788

<210> 9

<211> 595

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Polypeptide sequence for alphaprepro signal
 sequence fused to human P501S and fused to a His
 tag

35

<400> 9

Met Ser Phe Leu Asn Phe Thr Ala Val Leu Phe Ala Ala Ser Ser Ala
 1 5 10 15
 40 Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln Ile
 20 25 30
 Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe Asp
 35 40 45
 Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe
 50 55 60
 Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser
 65 70 75 80
 Leu Glu Lys Arg Glu Ala Glu Ala Met Val Leu Gly Ile Gly Pro Val
 85 90 95
 50 Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp
 100 105 110
 Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu
 115 120 125
 Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala
 130 135 140
 55 Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile
 145 150 155 160
 Leu Gly Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro
 165 170 175
 60 Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg

				180				185				190				
	Gln	Ala	Tyr	Ser	Val	Tyr	Ala	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu
			195					200					205			
5	Gly	Tyr	Leu	Leu	Pro	Ala	Ile	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro
		210					215					220				
	Tyr	Leu	Gly	Thr	Gln	Glu	Glu	Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile
	225					230					235				240	
	Phe	Leu	Thr	Cys	Val	Ala	Ala	Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala
				245							250				255	
10	Leu	Gly	Pro	Thr	Glu	Pro	Ala	Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Ser
				260					265					270		
	Pro	His	Cys	Cys	Pro	Cys	Arg	Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly
			275					280					285			
	Ala	Leu	Leu	Pro	Arg	Leu	His	Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr
15		290					295					300				
	Leu	Arg	Arg	Leu	Phe	Val	Ala	Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met
	305					310					315				320	
	Thr	Phe	Thr	Leu	Phe	Tyr	Thr	Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln
				325						330					335	
20	Gly	Val	Pro	Arg	Ala	Glu	Pro	Gly	Thr	Glu	Ala	Arg	Arg	His	Tyr	Asp
				340					345					350		
	Glu	Gly	Val	Arg	Met	Gly	Ser	Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile
			355					360					365			
25	Ser	Leu	Val	Phe	Ser	Leu	Val	Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly
		370					375					380				
	Thr	Arg	Ala	Val	Tyr	Leu	Ala	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala
	385					390				395					400	
	Gly	Ala	Thr	Cys	Leu	Ser	His	Ser	Val	Ala	Val	Val	Thr	Ala	Ser	Ala
				405						410					415	
30	Ala	Leu	Thr	Gly	Phe	Thr	Phe	Ser	Ala	Leu	Gln	Ile	Leu	Pro	Tyr	Thr
			420						425					430		
	Leu	Ala	Ser	Leu	Tyr	His	Arg	Glu	Lys	Gln	Val	Phe	Leu	Pro	Lys	Tyr
		435						440					445			
	Arg	Gly	Asp	Thr	Gly	Gly	Ala	Ser	Ser	Glu	Asp	Ser	Leu	Met	Thr	Ser
35		450					455					460				
	Phe	Leu	Pro	Gly	Pro	Lys	Pro	Gly	Ala	Pro	Phe	Pro	Asn	Gly	His	Val
	465					470					475				480	
	Gly	Ala	Gly	Gly	Ser	Gly	Leu	Leu	Pro	Pro	Pro	Pro	Ala	Leu	Cys	Gly
				485						490					495	
40	Ala	Ser	Ala	Cys	Asp	Val	Ser	Val	Arg	Val	Val	Val	Gly	Glu	Pro	Thr
				500					505					510		
	Glu	Ala	Arg	Val	Val	Pro	Gly	Arg	Gly	Ile	Cys	Leu	Asp	Leu	Ala	Ile
		515						520					525			
	Leu	Asp	Ser	Ala	Phe	Leu	Leu	Ser	Gln	Val	Ala	Pro	Ser	Leu	Phe	Met
45		530					535					540				
	Gly	Ser	Ile	Val	Gln	Leu	Ser	Gln	Ser	Val	Thr	Ala	Tyr	Met	Val	Ser
	545					550					555				560	
	Ala	Ala	Gly	Leu	Gly	Leu	Val	Ala	Ile	Tyr	Phe	Ala	Thr	Gln	Val	Val
				565						570					575	
50	Phe	Asp	Lys	Ser	Asp	Leu	Ala	Lys	Tyr	Ser	Ala	Gly	Gly	His	His	His
				580					585					590		
	His	His	His													
				595												

55

<210> 10

<211> 553

<212> PRT

<213> Mus musculus

60

<400> 10
Met Ile Gln Arg Leu Trp Ala Ser Arg Leu Leu Arg His Arg Lys Ala
1 5 10 15
5 Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
20 25 30
Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
35 40 45
Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
50 55 60
10 Leu Val Ser Val Pro Leu Leu Gly Ser Ala Ser Asp Gln Trp Arg Gly
65 70 75 80
Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Val
85 90 95
15 Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu
100 105 110
Leu Tyr Pro Asp Thr Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly
115 120 125
Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu
130 135 140
20 Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala
145 150 155 160
Phe Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr
165 170 175
25 Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Val Leu Ala Pro Tyr Leu
180 185 190
Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
195 200 205
Ile Cys Met Ala Ala Thr Leu Phe Val Thr Glu Glu Ala Val Leu Gly
210 215 220
30 Pro Pro Glu Pro Ala Glu Gly Leu Leu Val Ser Ala Val Ser Arg Arg
225 230 235 240
Cys Cys Pro Cys His Val Gly Leu Ala Phe Arg Asn Leu Gly Thr Leu
245 250 255
35 Phe Pro Arg Leu Gln Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
260 265 270
Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
275 280 285
Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
290 295 300
40 Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
305 310 315 320
Ile Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
325 330 335
45 Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Lys Phe Gly Thr Arg
340 345 350
Ser Val Tyr Leu Ala Ser Val Met Thr Phe Pro Val Ala Ala Ala Ala
355 360 365
Thr Cys Leu Ser His Ser Val Val Val Thr Ala Ser Ala Ala Leu
370 375 380
50 Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
385 390 395 400
Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
405 410 415
55 Asp Ala Gly Gly Ser Ser Gly Glu Asp Ser Gln Thr Thr Ser Phe Leu
420 425 430
Pro Gly Pro Lys Pro Gly Ala Leu Phe Pro Asn Gly His Val Gly Ser
435 440 445
Gly Ser Ser Gly Ile Leu Ala Pro Pro Pro Ala Leu Cys Gly Ala Ser
450 455 460
60 Ala Cys Asp Val Ser Met Arg Val Val Val Gly Glu Pro Pro Glu Ala

	465		470		475		480
	Arg Val Val Thr Gly	Arg Gly Ile Cys Leu Asp	Leu Ala Ile Leu Asp				
		485		490		495	
5	Ser Ala Phe Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser						
		500		505		510	
	Ile Val Gln Leu Ser His Ser Val Thr Ala Tyr Met Val Ser Ala Ala						
		515		520		525	
	Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp						
		530		535		540	
10	Lys Asn Asp Leu Ala Lys Tyr Ser Val						
	545		550				
	<210> 11						
15	<211> 2188						
	<212> DNA						
	<213> Mus musculus						
	<400> 11						
20	gagattttaa aggcgccccg tggcgcgcggt tgggtgamsca ggbgtcgccg agctcgcacg 60						
	cgccagcccc aggtgacagc cgcacgccgg gccaggatct gaccgacgag atgtgtcccc 120						
	atcaagcaag gcactagatg gtgacgtgtt tagcgtggga cgagatgctg aattggcact 180						
	aaagggctgg cagaaatggg aacctggctg caccctagga ggtagtgct agtgaggagg 240						
	agaagccacg gcagggctga ctcaaagcag ctgtggagta tgtgagtagc ccctggaacc 300						
25	ctacctgccc tgtccatcat gatccagagg ctgtgggcca gccgtctgct acggcaccgg 360						
	aaagctcagc tcctgctggt caacctgctc accttggcc tggaggtgtg cctggctgcc 420						
	ggcattacct atgtgccacc ccttctgctg gaagtcgggg tggaggagaa attcatgacc 480						
	atgggtgttg gcattggccc agtgctaggc ctggtttctg ttccactcct aggcctcagc 540						
	agtgaccagt ggcgtgggcg ctatggccgc cggagacct ttatctgggc tttgtccctg 600						
30	ggtgtcctgc taagcctctt tctcatcccg agggctggct ggctggcagg actgctgtac 660						
	ccagacacca ggcacctgga gttggccctg ctgactctgg gactggggct gctggacttt 720						
	tgtggccagg tgtgctttac tccattggag gccttactct ccgacctctt ccgggaccca 780						
	gaccactgcc gccaaagcct ctctgtctac gccttcatga tcagccttgg gggctgcctg 840						
	ggctacctct tacctgccat tgactgggac accagcgttc tggcccccta cctgggtact 900						
35	caggaagaat gcctctttgg cctcctcacc ctcatcttcc tcatctgcat ggcagccact 960						
	ctgtttgtga cggaggaggc agtactgggc ccacccgagc cggcagaagg gttgttggtc 1020						
	tctgccgtgt cgcgccgatg ctgcccatgc cacgttggcc tggctttccg gaatctgggt 1080						
	accctgtttc cccggctgca gcagctgtgc tgccgcagtc ctgcaccct acgccgactc 1140						
	tttgtggctg agctgtgcag ctggatggca cttatgactt tcacactgtt ctacacggac 1200						
40	ttcgtgggag aggggctgta ccagggtgta ccagagccg agccaggcac cgaggcccgg 1260						
	agacactatg atgaaggcat tccaatgggc agcctggggc tcttcctgca gtgtgccatc 1320						
	tccctggtct tctccctggt catggacagg ctggtacaga agttcggcac acggtcagtc 1380						
	tatctggcca gtgtgatgac ctttctctgt gctgccgtgc ccacctgcct gtcccacagc 1440						
	gtggtggtag tgacagcctc agctgcctc accgggttca ccttctcggc cttgcagatc 1500						
45	ctgccttaca cgctcgctc cctctaccac cgtgagaagc aggtgttcct gcccaaatac 1560						
	cgaggggacg ctggaggtag cagcggtag gacagccaga caaccagctt cttgccaggc 1620						
	cctaagccag gagctctctt ccccaatgga cacgtgggct ctggcagcag cggcatcctg 1680						
	gcccctccac ctgcactctg tggggcctct gcctgcgatg tctccatgcg agtgggtggtg 1740						
	ggtgagccac ctgaggccag ggttggttac ggacggggca tttgcctgga cctggccatt 1800						
50	ctggacagtg cctttctgct gtcccagggt gctccgtccc tgttcatggg ctccattgtc 1860						
	cagctagggc actctgtcac tgcctatatg gtatcagctg caggcttggg tctggctgcc 1920						
	atttactttg ctacacaggt agtgtttgac aagaacgact tggccaaata ctcagtgtag 1980						
	aattgtgtaa ggcacaaag agagggtctg cctcatgggt tctcagccct tagagggtctg 2040						
	cagagctggc ctctccaggt ctttgtcgcc taagtggctc tctgctgcca ccctaaggca 2100						
55	gtgaggtgta ttgttgaca gataggagcc agagctttcg gggctctggc ttcagagtct 2160						
	ggctggccta ctggcagcct ctgcgatg 2188						